#16 IDS 09/966,880

Report of EP 00911431.5
Your Ref.: JP-101DP2PCT-EP

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/52, 9/00, A61K 38/43, C07K 16/40, C12Q 1/68

A1

(11) International Publication Number: **WO 98/40494**

(43) International Publication Date: 17 September 1998 (17.09.98)

(21) International Application Number: PCT/US98/04565

(22) International Filing Date: 9 March 1998 (09.03.98)

(30) Priority Data:
08/816,241 13 March 1997 (13.03.97) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 08/816,241 (CIP)
Filed on 13 March 1997 (13.03.97)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). GOLI, Surya, K. [IN/US]; 3311 Cantamar Court, San Jose, CA 95135 (US).

(74) Agent: BILLINGS, Lucy, J.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

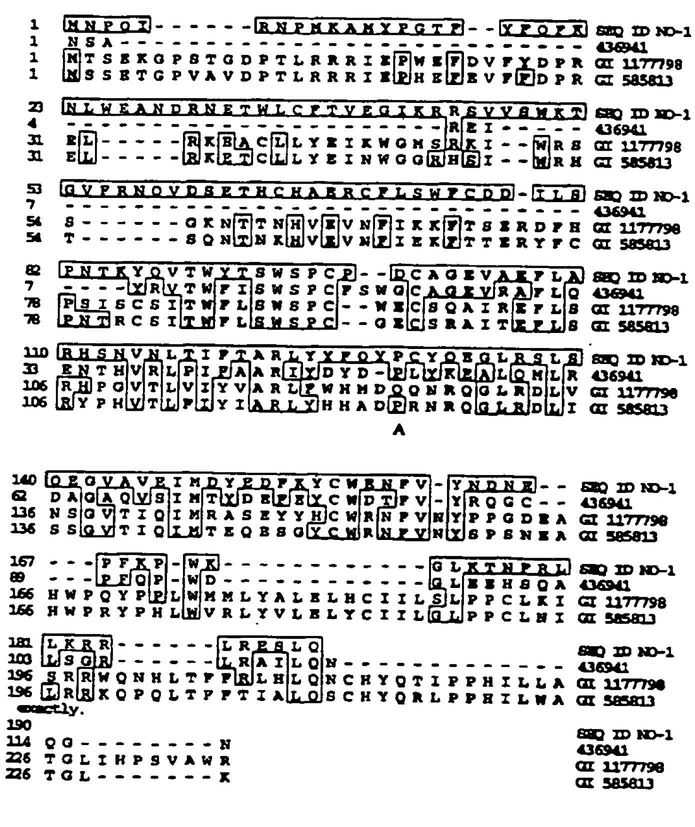
(81) Designated States: AT, AU, BR, CA, CH, CN, DE, DK, ES, FI, GB, IL, JP, KR, MX, NO, NZ, RU, SE, SG, US, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: NOVEL RNA EDITING ENZYME

(57) Abstract

The present invention provides a human RNA editing enzyme (REE-2) and polynucleotides which identify and encode REE-2. The invention also provides expression vectors and host cells, agonists, antibodies or antagonists. The invention provides methods for producing REE-2 and for treating diseases associated with expression of REE-2.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

NOVEL RNA EDITING ENZYME

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of a novel RNA editing enzyme and to the use of these sequences in the diagnosis, prevention, and treatment of cancer, autoimmune disorders, circulatory system disorders, viral diseases, and neurological diseases.

BACKGROUND ART

Apolipoprotein B (apo B) circulates in two distinct forms referred to as apo B100 and apo B48. Apo B48 is encoded by same gene as apo B100 and arises as a result of mRNA editing. A single cytidine nucleotide in apo B100 mRNA is deaminated by a zinc-containing enzyme resulting in a CAA to UAA-stop codon change (Navaratnam N. et al. (1995) Cell 81:187-195). RNA editing of apo B has important consequences in the catabolism of plasma lipoproteins and the ability to generate hybrid lipoproteins (Davidson N.O. (1993) Ann. Med. 25:539-53).

Hadjiagapiou C. et al. (1994, Nucleic Acids Res. 22:1874-1879) cloned the apo B mRNA editing protein (HEPR) from human small intestine cDNA. HEPR is the catalytic subunit of an enzyme complex which includes, yet to be identified, complementation factors (Teng B. et al. (1993) Science 260:1819). HEPR contains consensus phosphorylation sites as well as conserved histidine and cysteine residues identified as a Zn^{2+} binding motif in other cytidine deaminases. Mutational studies indicated that the putative zinc-coordinating residues H_{61}, C_{93}, and C_{96} and the catalytically active residues E_{63} and P_{92} are necessary for both RNA editing and cytidine deaminase activities (MacGinnitie et al. (1995) J. Biol. Chem. 270:14768-14775). H_{61} is also required for RNA binding activity.

HEPR is expressed in the adult small intestine and to a much lesser extent in the stomach, colon, and testis (Hadjiagapiou et al, supra). The rabbit homolog of HEPR is only expressed in the small and large intestine, and the complementation proteins, essential for RNA editing, were found to exist in a wide range of organs that do not express apo B. This suggests the presence of additional RNA editing enzymes with a more widespread role in the generation of RNA and protein diversity (Yamanaka S. et al. (1994) J. Biol. Chem. 269:21725-21734; Hodges P. et al. (1992) Trends Biochem. Sci. 17:77-81).

The principle of therapeutic RNA editing was demonstrated by using cell extracts containing a RNA editing enzyme to correct an aberrant stop codon introduced into synthetic RNA encoding dystrophin protein (Woolf, T.M. et al. (1995) Proc. Natl. Acad. Sci. 92:8298-8302). Deamination of a nucleotide in a stop codon resulted in translation read-through and a dramatic increase in expression of a downstream gene.

BNSDOCID: <WO___9840494A1_I_>

Apo B editing is a mechanism which determines how much apo B48 is synthesized in place of apo B100 in a specific tissue. Apo B100 is the exclusive apolipoprotein of low density lipoproteins which transport most of the plasma cholesterol in humans. In contrast, apo B48 is directed to chylomicrons, triglyceride-rich lipoproteins that transport dietary lipids and undergo catabolic clearance much faster than particles containing apo B100 (Young, S.G. (1990) Circulation 82:1574-1594). Apo B editing has major physiological and clinical implications. All apo B100 containing lipoproteins are atherogenic, especially when present in high concentrations. Alterations in apo B100 can cause either hypocholesterolemia or hypercholesterolemia (Innerarity, T.L. et al. (1991) Adv. Exp. Med. Biol. 285:25-31). Apo B mRNA editing decreases the amount of apo B100 production. Somatic gene transfer of rat REPR, the rat homolog of HEPR, into the liver of mice essentially eliminates apo B100 and plasma low-density lipoprotein without affecting anti-atherogenic high density lipoproteins (Teng, B. et al. (1994) J. Biol. Chem. 269:29395-29404).

Alpha-galactosidase is a lysosomal enzyme that is deficient in patients with Fabry's disease. After excluding other possibilities, Novo, F.J. et al. (1995, Nucleic Acids Res. 23:2636-2640) proposed that RNA editing is responsible for an observed nucleotide conversion in alpha-galactosidase mRNA from Fabry's disease patients. The enzyme responsible for the nucleotide conversion has not been identified.

Among the main characteristics of psoriasis are abnormal keratinocyte proliferation and differentiation. A search for proteins that are implicated in psoriasis yielded a partial sequence for phorbolin I, a protein upregulated in psoriatic keratinocytes (Madsen, P.P., unpublished).

Deamination of RNA nucleotides also occurs in the brain where RED1, a double stranded RNA adenosine deaminase edits the channel determinant site in glutamate receptor pre-mRNA. This site controls the Ca^{2+} permeability of alpha-amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA) receptors. Another double stranded RNA adenosine deaminase, DRADA, edits a different site in the glutamate receptor pre-mRNA (Kim, U. et al. (1994) J. Biol. Chem. 269: 13480-13489). Glutamate receptor RNA editing is differentially regulated (Nutt, S.L. (1994) Neuroreport 5: 1679-1683). Since glutamate receptors are essential for fast excitatory neurotransmission, RNA editing may play a critical role in normal brain function and development. Furthermore, dysfunction of RNA editing may have neuropathological consequences and could be related to neurodegenerative diseases (Nutt, et al, supra). Evidence suggests that RED1 and DRADA are members of a larger gene family of enzymes that deaminate nuclear transcripts and have distinct but overlapping substrate specificities (Melcher, T. et al.

BNSDOCID <WO___9840494A1_I_>

(1996) Nature 379:460-464).

Neurofibromatosis type I is a common neurological disease in which tissue derived from the embryonic neural crest is predisposed to the development of cancer. The NF1 gene encodes a tumor suppressor that is responsible for neurofibromatosis type I. NF1 mRNA undergoes C to U editing, creating an in-frame stop codon, in a manner similar to apoB mRNA editing. However, evidence suggests that the apoB complex does not function in NF1 mRNA editing (Skuse, G. R. et al. (1996) Nuc. Acids Res. 24:478-485). Cappione, A. J. et al. found that unusually high levels of NF1 mRNA editing occur in tumor tissues associated with neurofibromatosis type I (1997; Am. J. Hum. Genet. 60:305-312).

Discovery of proteins related to human RNA editing enzymes and the polynucleotides encoding them satisfies a need in the art by providing new compositions useful in diagnosis, prevention, and treatment of cancer, autoimmune disorders, circulatory system disorders, viral diseases, and neurological diseases.

DISCLOSURE OF THE INVENTION

The present invention features a novel human RNA editing enzyme hereinafter designated REE-2 and characterized as having similarity to the mRNA editing enzymes phorbolin I, HEPR, and REPR.

Accordingly, the invention features a substantially purified REE-2 having the amino acid sequence shown in SEQ ID NO:1.

One aspect of the invention features isolated and substantially purified polynucleotides that encode REE-2. In a particular aspect, the polynucleotide is the nucleotide sequence of SEQ ID NO:2.

The invention also relates to a polynucleotide sequence comprising the complement of SEQ ID NO:2 or variants thereof. In addition, the invention features polynucleotide sequences which hybridize under stringent conditions to SEQ ID NO:2.

The invention additionally features nucleic acid sequences encoding polypeptides, oligonucleotides, peptide nucleic acids (PNA), fragments, portions or antisense molecules thereof, and expression vectors and host cells comprising polynucleotides that encode REE-2. The present invention also features antibodies which bind specifically to REE-2, and pharmaceutical compositions comprising substantially purified REE-2. The invention also features agonists and antagonists of REE-2. The invention also features a method for producing REE-2, and methods for treating cancer, circulatory system disorders, viral diseases, and neurological diseases by administering REE-2 or an REE-2 fragment or derivative thereof.

BNSDOCID: <WO___9840494A1_I_>

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A and 1B show the amino acid sequence (SEQ ID NO:1) and nucleic acid sequence (SEQ ID NO:2) of REE-2. The alignment was produced using MacDNASIS PRO™ software (Hitachi Software Engineering Co., Ltd., San Bruno, CA).

Figures 2A and 2B show the amino acid sequence alignments among REE-2 (SEQ ID NO:1), phorbolin I (GI 436941; SEQ ID NO:3), HEPR (GI 1177798; SEQ ID NO:4), and REPR (GI 585813; SEQ ID NO:5). The alignment was produced using the multisequence alignment program of DNASTAR™ software (DNASTAR Inc, Madison WI).

Figure 3 shows the hydrophobicity plot (MacDNASIS PRO software) for REE-2, SEQ ID NO: 1; the positive X axis reflects amino acid position, and the negative Y axis, hydrophobicity.

Figure 4 shows the hydrophobicity plot for HEPR, SEQ ID NO:4.

MODES FOR CARRYING OUT THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors, and reagents described as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a host cell" includes a plurality of such host cells, reference to the "antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods, devices, and materials are now described. All publications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the cell lines, vectors, and methodologies which are reported in the publications which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide, or

BNSDOCID: <WO__9840494A1_I_>

polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand. Similarly, "amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules.

Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms, such as "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

"Peptide nucleic acid", as used herein, refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary strand of nucleic acid (Nielsen, P.E. et al. (1993) Anticancer Drug Des. 8:53-63).

REE-2, as used herein, refers to the amino acid sequences of substantially purified REE-2 obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source whether natural, synthetic, semi-synthetic, or recombinant.

"Consensus", as used herein, refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, or which has been extended using XL-PCR™ (Perkin Elmer, Norwalk, CT) in the 5' and/or the 3' direction and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte clone using the GELVIEW™ Fragment Assembly system (GCG, Madison, WI), or which has been both extended and assembled.

A "variant" of REE-2, as used herein, refers to an amino acid sequence that is altered by one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More rarely, a variant may have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNASTAR software.

A "deletion", as used herein, refers to a change in either amino acid or nucleotide sequence in which one or more amino acid or nucleotide residues, respectively, are absent.

BNSDOCID <WO___9840494A1_I_>

An "insertion" or "addition", as used herein, refers to a change in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid or nucleotide residues, respectively, as compared to the naturally occurring molecule.

A "substitution", as used herein, refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

The term "biologically active", as used herein, refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic REE-2, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "agonist", as used herein, refers to a molecule which, when bound to REE-2, causes a change in REE-2 which modulates the activity of REE-2. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to REE-2.

The terms "antagonist" or "inhibitor", as used herein, refer to a molecule which, when bound to REE-2, blocks or modulates the biological or immunological activity of REE-2. Antagonists and inhibitors may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to REE-2.

The term "modulate", as used herein, refers to a change or an alteration in the biological activity of REE-2. Modulation may be an increase or a decrease in protein activity, a change in binding characteristics, or any other change in the biological, functional or immunological properties of REE-2.

The term "mimetic", as used herein, refers to a molecule, the structure of which is developed from knowledge of the structure of REE-2 or portions thereof and, as such, is able to effect some or all of the actions of RNA editing enzyme-like molecules.

The term "derivative", as used herein, refers to the chemical modification of a nucleic acid encoding REE-2 or the encoded REE-2. Illustrative of such modifications would be replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative would encode a polypeptide which retains essential biological characteristics of the natural molecule.

The term "substantially purified", as used herein, refers to nucleic or amino acid sequences that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated.

"Amplification" as used herein refers to the production of additional copies of a nucleic

BNSDOCID: <WO__9840494A1_I_>

acid sequence and is generally carried out using polymerase chain reaction (PCR) technologies well known in the art (Dieffenbach, C.W. and G.S. Dveksler (1995) PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, NY).

The term "hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex", as used herein, refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen binds between complementary G and C bases and between complementary A and T bases; these hydrogen bonds may be further stabilized by base stacking interactions. The two complementary nucleic acid sequences hydrogen bond in an antiparallel configuration. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., membranes, filters, chips, pins or glass slides to which cells have been fixed for in situ hybridization).

The terms "complementary" or "complementarity", as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, for the sequence "A-G-T" binds to the complementary sequence "T-C-A". Complementarity between two single-stranded molecules may be "partial", in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands.

The term "homology", as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence is one that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid; it is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The

BNSDOCID <WO__9840494A1_I_>

absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% identity); in the absence of non-specific binding, the probe will not hybridize to the second non-complementary target sequence.

As known in the art, numerous equivalent conditions may be employed to comprise either low or high stringency conditions. Factors such as the length and nature (DNA, RNA, base composition) of the sequence, nature of the target (DNA, RNA, base composition, presence in solution or immobilization, etc.), and the concentration of the salts and other components (e.g., the presence or absence of formamide, dextran sulfate and/or polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of either low or high stringency different from, but equivalent to, the above listed conditions.

The term "stringent conditions", as used herein, is the "stringency" which occurs within a range from about Tm-5°C (5°C below the melting temperature (Tm) of the probe) to about 20°C to 25°C below Tm. As will be understood by those of skill in the art, the stringency of hybridization may be altered in order to identify or detect identical or related polynucleotide sequences.

The term "antisense", as used herein, refers to nucleotide sequences which are complementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of interest in a reverse orientation to a viral promoter which permits the synthesis of a complementary strand. Once introduced into a cell, this transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further transcription or translation. In this manner, mutant phenotypes may be generated. The designation "negative" is sometimes used in reference to the antisense strand, and "positive" is sometimes used in reference to the sense strand.

The term "portion", as used herein, with regard to a protein (as in "a portion of a given protein") refers to fragments of that protein. The fragments may range in size from four amino acid residues to the entire amino acid sequence minus one amino acid. Thus, a protein "comprising at least a portion of the amino acid sequence of SEQ ID NO:1" encompasses the full-length human REE-2 and fragments thereof.

"Transformation", as defined herein, describes a process by which exogenous DNA enters and changes a recipient cell. It may occur under natural or artificial conditions using various

BNSDOCID: <WO___9840494A1_I_>

methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the host cell being transformed and may include, but is not limited to, viral infection, electroporation, lipofection, and particle bombardment. Such "transformed" cells include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. They also include cells which transiently express the inserted DNA or RNA for limited periods of time.

The term "antigenic determinant", as used herein, refers to that portion of a molecule that makes contact with a particular antibody (i.e., an epitope). When a protein or fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to a given region or three-dimensional structure on the protein; these regions or structures are referred to as antigenic determinants. An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The terms "specific binding" or "specifically binding", as used herein, in reference to the interaction of an antibody and a protein or peptide, mean that the interaction is dependent upon the presence of a particular structure (i.e., the antigenic determinant or epitope) on the protein; in other words, the antibody is recognizing and binding to a specific protein structure rather than to proteins in general. For example, if an antibody is specific for epitope "A", the presence of a protein containing epitope A (or free, unlabeled A) in a reaction containing labeled "A" and the antibody will reduce the amount of labeled A bound to the antibody.

The term "sample", as used herein, is used in its broadest sense. A biological sample suspected of containing nucleic acid encoding REE-2 or fragments thereof may comprise a cell, chromosomes isolated from a cell (e.g., a spread of metaphase chromosomes), genomic DNA (in solution or bound to a solid support such as for Southern analysis), RNA (in solution or bound to a solid support such as for northern analysis), cDNA (in solution or bound to a solid support), an extract from cells or a tissue, and the like.

The term "correlates with expression of a polynucleotide", as used herein, indicates that the detection of the presence of ribonucleic acid that is similar to SEQ ID NO:2 by northern analysis is indicative of the presence of mRNA encoding REE-2 in a sample and thereby correlates with expression of the transcript from the polynucleotide encoding the protein.

"Alterations" in the polynucleotide of SEQ ID NO: 2, as used herein, comprise any alteration in the sequence of polynucleotides encoding REE-2 including deletions, insertions, and

BNSDOCID <WO__9840494A1_I_>

point mutations that may be detected using hybridization assays. Included within this definition is the detection of alterations to the genomic DNA sequence which encodes REE-2 (e.g., by alterations in the pattern of restriction fragment length polymorphisms capable of hybridizing to SEQ ID NO:2), the inability of a selected fragment of SEQ ID NO: 2 to hybridize to a sample of genomic DNA (e.g., using allele-specific oligonucleotide probes), and improper or unexpected hybridization, such as hybridization to a locus other than the normal chromosomal locus for the polynucleotide sequence encoding REE-2 (e.g., using fluorescent in situ hybridization [FISH] to metaphase chromosomes spreads).

As used herein, the term "antibody" refers to intact molecules as well as fragments thereof, such as Fa, $F(ab')_2$, and Fv, which are capable of binding the epitopic determinant. Antibodies that bind REE-2 polypeptides can be prepared using intact polypeptides or fragments containing small peptides of interest as the immunizing antigen. The polypeptide or peptide used to immunize an animal can be derived from the transition of RNA or synthesized chemically, and can be conjugated to a carrier protein, if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin and thyroglobulin. The coupled peptide is then used to immunize the animal (e.g., a mouse, a rat, or a rabbit).

The term "humanized antibody", as used herein, refers to antibody molecules in which amino acids have been replaced in the non-antigen binding regions in order to more closely resemble a human antibody, while still retaining the original binding ability.

THE INVENTION

The invention is based on the discovery of a novel human RNA editing enzyme, (REE-2), the polynucleotides encoding REE-2, and the use of these compositions for the diagnosis, prevention, or treatment of cancer, autoimmune disorders, circulatory system disorders, viral diseases, and neurological diseases.

Nucleic acids encoding the human REE-2 of the present invention were first identified in Incyte Clone 1646823 from the prostate tumor tissue cDNA library (PROSTUT09) through a computer-generated search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:2, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 1646823 (PROSTUT09), 1354426 (LUNGNOT09), and 1681742 (STOMFET01).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1, as shown in Figures 1A and 1B. REE-2 is 190 amino acids in length and has chemical and structural homology with phorbolin I (GI 436941; SEQ ID NO:3), HEPR (GI 1177798; SEQ ID NO:4), and REPR (GI 585813; SEQ ID NO:5; Figures 2A and 2B). In

BNSDOCID: <WO_____9840494A1_I_>

particular, REE-2 shares 28% identity with HEPR, 30% identity with REPR, and 43% identity with a partial sequence of phorbolin I. REE-2 contains conserved zinc-coordinating residues C_{97} and C_{100}, and catalytically active residues E_{68} and P_{96}, necessary for both RNA editing and cytidine deaminase activities. REE also contains conserved residue H_{66}, required for RNA binding, RNA editing, and cytidine deaminase activities. As illustrated by Figures 3 and 4, REE-2 and HEPR have rather similar hydrophobicity plots. Northern analysis revealed the expression of this sequence in various cDNA libraries, a high proportion of which are derived from tumors, neuronal tissues, immune system cells, or synovial tissue from arthritis patients.

The invention also encompasses REE-2 variants. A preferred REE-2 variant is one having at least 80%, and more preferably 90%, amino acid sequence similarity to the REE-2 amino acid sequence (SEQ ID NO:1). A most preferred REE-2 variant is one having at least 95% amino acid sequence similarity to SEQ ID NO:1.

The invention also encompasses polynucleotides which encode REE-2. Accordingly, any nucleic acid sequence which encodes the amino acid sequence of REE-2 can be used to generate recombinant molecules which express REE-2. In a particular embodiment, the invention encompasses the polynucleotide comprising the nucleic acid sequence of SEQ ID NO:2 as shown in Figures 1A and 1B.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding REE-2, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring REE-2, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode REE-2 and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring REE-2 under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding REE-2 or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding REE-2 and its derivatives without altering the encoded amino acid sequences include

BNSDOCID: <WO___9840494A1_I_>

the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences, or portions thereof, which encode REE-2 and its derivatives, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents that are well known in the art at the time of the filing of this application. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding REE-2 or any portion thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those shown in SEQ ID NO:2, under various conditions of stringency. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Wahl, G.M. and S.L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A.R. (1987; Methods Enzymol. 152:507-511), and may be used at a defined stringency.

Altered nucleic acid sequences encoding REE-2 which are encompassed by the invention include deletions, insertions, or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent REE-2. The encoded protein may also contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent REE-2. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of REE-2 is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid; positively charged amino acids may include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; phenylalanine and tyrosine.

Also included within the scope of the present invention are alleles of the genes encoding REE-2. As used herein, an "allele" or "allelic sequence" is an alternative form of the gene which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one, or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

BNSDOCID <WO___9840494A1_I_>

Methods for DNA sequencing which are well known and generally available in the art may be used to practice any embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase® (US Biochemical Corp, Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham, Chicago, IL), or combinations of recombinant polymerases and proofreading exonucleases such as the ELONGASE Amplification System marketed by Gibco BRL (Gaithersburg, MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) and the ABI 377 DNA sequencers (Perkin Elmer).

The nucleic acid sequences encoding REE-2 may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed, "restriction-site" PCR, uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). In particular, genomic DNA is first amplified in the presence of primer to linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR may also be used to amplify or extend sequences using divergent primers based on a known region (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). The primers may be designed using OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, MN), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may also be used to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before performing PCR.

Another method which may be used to retrieve unknown sequences is that of Parker, J.D.

BNSDOCID <WO__9840494A1_I_>

et al. (1991; Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PromoterFinder™ libraries to walk in genomic DNA (Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into the 5' and 3' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled devise camera. Output/light intensity may be converted to electrical signal using appropriate software (e.g. Genotyper™ and Sequence Navigator™, Perkin Elmer) and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode REE-2, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of REE-2 in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express REE-2.

As will be understood by those of skill in the art, it may be advantageous to produce REE-2-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter REE-2 encoding sequences for a variety of reasons,

BNSDOCID: <WO___9840494A1_I_>

including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding REE-2 may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of REE-2 activity, it may be useful to encode a chimeric REE-2 protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the REE-2 encoding sequence and the heterologous protein sequence, so that REE-2 may be cleaved and purified away from the heterologous moiety.

In another embodiment, sequences encoding REE-2 may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of REE-2, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J.Y. et al. (1995) Science 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer).

The newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, NY). The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; Creighton, supra). Additionally, the amino acid sequence of REE-2, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a biologically active REE-2, the nucleotide sequences encoding REE-2 or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence.

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding REE-2 and appropriate transcriptional and

BNSDOCID <WO__9840494A1_I_>

translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, NY, and Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY.

A variety of expression vector/host systems may be utilized to contain and express sequences encoding REE-2. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript® phagemid (Stratagene, LaJolla, CA) or pSport1™ plasmid (Gibco BRL) and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (e.g., heat shock, RUBISCO; and storage protein genes) or from plant viruses (e.g., viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding REE-2, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for REE-2. For example, when large quantities of REE-2 are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional E. coli cloning and expression vectors such as Bluescript® (Stratagene), in which the sequence encoding REE-2 may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of ß-galactosidase so that a hybrid protein is produced; pIN

BNSDOCID <WO___9840494A1_I_>

vectors (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors (Promega, Madison, WI) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) Methods Enzymol. 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding REE-2 may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L.E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, NY; pp. 191-196.

An insect system may also be used to express REE-2. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The sequences encoding REE-2 may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of REE-2 will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, S. frugiperda cells or Trichoplusia larvae in which REE-2 may be expressed (Engelhard, E.K. et al. (1994) Proc. Nat. Acad. Sci. 91:3224-3227).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding REE-2 may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and

BNSDOCID: <WO___9840494A1_I_>

tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing REE-2 in infected host cells (Logan, J. and Shenk, T. (1984) Proc. Natl. Acad. Sci. 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding REE-2. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding REE-2, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express REE-2 may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using

BNSDOCID <WO___9840494A1_I_>

tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1980) Cell 22:817-23) genes which can be employed in tk^- or $aprt^-$ cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, supra). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, ß glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. et al. (1995) Methods Mol. Biol. 55:121-131). .

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding REE-2 is inserted within a marker gene sequence, recombinant cells containing sequences encoding REE-2 can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding REE-2 under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain the nucleic acid sequence encoding REE-2 and express REE-2 may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

The presence of polynucleotide sequences encoding REE-2 can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or portions or fragments of polynucleotides encoding REE-2. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences encoding REE-2 to detect transformants

BNSDOCID <WO__9840494A1_I_>

containing DNA or RNA encoding REE-2. As used herein "oligonucleotides" or "oligomers" refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.

A variety of protocols for detecting and measuring the expression of REE-2, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on REE-2 is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, MN) and Maddox, D.E. et al. (1983; J. Exp. Med. 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding REE-2 include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding REE-2, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, MI); Promega (Madison WI); and U.S. Biochemical Corp., Cleveland, OH). Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding REE-2 may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode REE-2 may be designed to contain signal sequences which direct secretion of REE-2 through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding REE-2 to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins.

BNSDOCID: <WO___9840494A1_I_>

Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, WA). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, CA) between the purification domain and REE-2 may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing REE-2 and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography as described in Porath, J. et al. (1992, Prot. Exp. Purif. 3:263-281) while the enterokinase cleavage site provides a means for purifying REE-2 from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D.J. et al. (1993; DNA Cell Biol. 12:441-453).

In addition to recombinant production, fragments of REE-2 may be produced by direct peptide synthesis using solid-phase techniques Merrifield J. (1963) J. Am. Chem. Soc. 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of REE-2 may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

THERAPEUTICS

Based on the chemical and structural homology among REE-2, phorbolin I, HEPR, and REPR; and the expression of REE-2 in tissue derived from cancer patients, REE-2 appears to be associated with the development of cancer, autoimmune disorders, circulatory system disorders, viral diseases, and neurological diseases.

Directed RNA editing may be used to alter transcripts of mutated genes that are responsible for cancerous growth. Therefore, in one embodiment, REE-2 or a fragment or derivative thereof may be administered to a subject to treat or prevent cancer. Types of cancer include, but are not limited to: adenocarcinoma; leukemia; melanoma; lymphoma; sarcoma; and cancers of the colon, liver, brain, small intestine, large intestine, breast, ovary, kidney, lung, and prostate.

In another embodiment, a vector capable of expressing REE-2, or a fragment or a derivative thereof, may also be administered to a subject to treat or prevent cancer including, but not limited, the types of cancers listed above.

BNSDOCID <WO__9840494A1_I_>

Directed RNA editing may be used to alter transcripts of viral genes that are responsible for infectious diseases. Therefore, in one embodiment, REE-2 or a fragment or derivative thereof may be administered to a subject to treat or prevent viral diseases. Types of viral diseases include, but are not limited to, those caused by HIV, hantavirus, poxyviruses, herpesviruses, papillomaviruses, polyomaviruses, adenoviruses, picornaviruses, and togaviruses.

In another embodiment, a vector capable of expressing REE-2, or a fragment or a derivative thereof, may also be administered to a subject to treat or prevent viral diseases including, but not limited, the types of viral diseases listed above.

REE-2 is similar to phorbolin I, a protein whose expression rises in the cells of patients suffering from an autoimmune system disorder. Therefore, in one embodiment, an antagonist to REE-2 may be administered to a subject to treat or prevent an autoimmune disorder. Such disorders include but are not limited to, rheumatoid arthritis, multiple sclerosis, scleroderma, psoriasis, Grave's disease, Sjogren's disease, Crohn's disease, diabetes, lupus, allergies, asthma, and myasthenia gravis.

When adminitered to a patient, REE-2 may edit the transcripts of Apo B100 or other lipid transport proteins. This in turn may prevent the accumulation of Apo B100 or other disease causing lipoproteins. Therefore, in one embodiment, REE-2 or a fragment or derivative thereof may be administered to a subject to treat or prevent a circulatory system disorder. Such disorders include but are not limited to, atherosclerosis, hypercholesterolemia, hypocholesterolemia, and stroke.

In another embodiment, a vector capable of expressing REE-2, or a fragment or a derivative thereof, may also be administered to a subject to treat or prevent a circulatory system disorder including, but not limited, the circulatory system disorders listed above.

A high level of RNA editing activity has been correlated with the pathogenesis of neurofibromatosis, a neurological disease. Therefore, in one embodiment, an antagonist to REE-2 may be administered to a subject to treat or prevent a neurological disease. Such neurological diseases include but are not limited to, neurofibromatosis, Alzheimer's disease, Parkinson's disease, neural tube defects, multiple sclerosis, dementia, and learning disabilities.

In other embodiments, any of the therapeutic proteins, antagonists, antibodies, agonists, antisense sequences or vectors described above may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment

BNSDOCID <WO__9840494A1_I_>

or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Antagonists or inhibitors of REE-2 may be produced using methods which are generally known in the art. In particular, purified REE-2 may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind REE-2.

The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with REE-2 or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the peptides, fragments, or oligopeptides used to induce antibodies to REE-2 have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids. It is also preferable that they are identical to a portion of the amino acid sequence of the natural protein, and they may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of REE-2 amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule.

Monoclonal antibodies to REE-2 may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen

BNSDOCID <WO__9840494A1_I_>

specificity and biological activity can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce REE-2-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:11120-3).

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for REE-2 may also be generated. For example, such fragments include, but are not limited to, the F(ab')2 fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 254:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between REE-2 and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering REE-2 epitopes is preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides encoding REE-2, or any fragment thereof, or antisense molecules, may be used for therapeutic purposes. In one aspect, antisense to the polynucleotide encoding REE-2 may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding REE-2. Thus, antisense molecules may be used to modulate REE-2 activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding REE-2.

BNSDOCID <WO___9840494A1_I_>

Expression vectors derived from retro viruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods which are well known to those skilled in the art can be used to construct recombinant vectors which will express antisense molecules complementary to the polynucleotides of the gene encoding REE-2. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra).

Genes encoding REE-2 can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide or fragment thereof which encodes REE-2. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, DNA, RNA, or PNA, to the control regions of the gene encoding REE-2, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples which may be used include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding REE-2.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be

BNSDOCID <WO__9840494A1_I_>

evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding REE-2. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods which are well known in the art.

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of REE-2, antibodies to REE-2, mimetics, agonists, antagonists, or inhibitors of REE-2. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical

BNSDOCID <WO__9840494A1_I_>

carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of

BNSDOCID <WO___9840494A1_I_>

gelatin. as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches. lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules. the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils. liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose. sorbitol. or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of REE-2, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

BNSDOCID <WO___9840494A1_I_>

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example REE-2 or fragments thereof, antibodies of REE-2, agonists, antagonists or inhibitors of REE-2, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind REE-2 may be used for the

BNSDOCID <WO__9840494A1_I_>

diagnosis of conditions or diseases characterized by expression of REE-2, or in assays to monitor patients being treated with REE-2, agonists, antagonists or inhibitors. The antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for REE-2 include methods which utilize the antibody and a label to detect REE-2 in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used, several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring REE-2 are known in the art and provide a basis for diagnosing altered or abnormal levels of REE-2 expression. Normal or standard values for REE-2 expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to REE-2 under conditions suitable for complex formation The amount of standard complex formation may be quantified by various methods, but preferably by photometric, means. Quantities of REE-2 expressed in subject, control and disease, samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding REE-2 may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of REE-2 may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess expression of REE-2, and to monitor regulation of REE-2 levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding REE-2 or closely related molecules, may be used to identify nucleic acid sequences which encode REE-2. The specificity of the probe, whether it is made from a highly specific region, e.g., 10 unique nucleotides in the 5' regulatory region, or a less specific region, e.g., especially in the 3' coding region, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding REE-2, alleles, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably

BNSDOCID <WO___9840494A1_I_>

contain at least 50% of the nucleotides from any of the REE-2 encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and derived from the nucleotide sequence of SEQ ID NO:2 or from genomic sequence including promoter, enhancer elements, and introns of the naturally occurring REE-2.

Means for producing specific hybridization probes for DNAs encoding REE-2 include the cloning of nucleic acid sequences encoding REE-2 or REE-2 derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as 32P or 35S, or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding REE-2 may be used for the diagnosis of conditions or diseases which are associated with expression of REE-2. Examples of such conditions or diseases include, but are not limited to, cancer of the prostate, bladder, pancreas, colon, thyroid, and brain; rheumatoid arthritis; and inflamed adenoids. The polynucleotide sequences encoding REE-2 may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; or in dip stick, pIN, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered REE-2 expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding REE-2 may be useful in assays that detect activation or induction of various cancers, particularly those mentioned above. The nucleotide sequences encoding REE-2 may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable control sample, the nucleotide sequences have hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding REE-2 in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of disease associated with expression of REE-

BNSDOCID <WO__9840494A1_I_>

2, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, which encodes REE-2, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease.

Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding REE-2 may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically, or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5'->3') and another with antisense (3'<-5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of REE-2 include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated (Melby, P.C. et al. (1993) J. Immunol. Methods, 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or

BNSDOCID: <WO___9840494A1_I_>

colorimetric response gives rapid quantitation.

In another embodiment of the invention, the nucleic acid sequences which encode REE-2 may also be used to generate hybridization probes which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques . Such techniques include FISH, FACS, or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C.M. (1993) Blood Rev. 7:127-134, and Trask, B.J. (1991) Trends Genet. 7:149-154.

FISH (as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, NY) may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene encoding REE-2 on a physical chromosomal map and a specific disease , or predisposition to a specific disease, may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier, or affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R.A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals.

In another embodiment of the invention, REE-2, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of

BNSDOCID: <WO___9840494A1_I_>

binding complexes, between REE-2 and the agent being tested, may be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to REE-2 large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with REE-2, or fragments thereof, and washed. Bound REE-2 is then detected by methods well known in the art. Purified REE-2 can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding REE-2 specifically compete with a test compound for binding REE-2. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with REE-2.

In additional embodiments, the nucleotide sequences which encode REE-2 may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

INDUSTRIAL APPLICABILITY

I PROSTUT09 cDNA Library Construction

The PROSTUT09 cDNA library was constructed from prostate tumor obtained from a 66-year-old Caucasian male. Surgery included a radical prostatectomy, a radical cystectomy, and a urinary diversion to the intestine. The pathology report indicated an invasive grade 3 (of 3) transitional cell carcinoma located within the prostatic urethra, and extended to involve periprostatic glands and diffusely invade the prostatic parencyma anteriorly and posteriorly. All final surgical margins including ureters (left and right, after multiple re-excisions) and prostatic urethra were negative for tumor. In addition to extensive involvement by transitional cell carcinoma, the right prostate contained a microscopic focus of adenocarcinoma, Gleason grade 3+2, which was confined to the prostate and showed no capsular penetration. Multiple right and left pelvic lymph nodes were negative for tumor. The patient was diagnosed with a malignant neoplasm of the prostate. The patient history included a previous transurethral prostatectomy and

BNSDOCID <WO__9840494A1_I_>

neoplasm of uncertain behavior of the lung. The patient history included benign hypertension, cerebrovascular disease, arteriosclerotic coronary artery disease, and tobacco use. The patient was taking insulin and Dyazide® (diuretic/antihypertensive; SmithKline Beecham Pharmaceuticals, Philadelphia, PA) at the time of surgery. The patient's family history included benign hypertension, cerebrovascular disease, and arteriosclerotic coronary artery disease in the patient's sibling, a malignant neoplasm of the breast in the patient's mother, a malignant neoplasm of the upper lobe of the lung in the patient's sibling, and a primary tuberculosis infection in the patient's father.

The frozen tissue was homogenized and lysed using a Brinkmann Homogenizer Polytron PT-3000 (Brinkmann Instruments, Westbury, NY) in guanidinium isothiocyanate solution. The lysate was centrifuged over a 5.7 M CsCl cushion using an Beckman SW28 rotor in a Beckman L8-70M Ultracentrifuge (Beckman Instruments) for 18 hours at 25,000 rpm at ambient temperature. The RNA was extracted with acid phenol pH 4.7, precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in RNAse-free water, and DNase treated at 37°C. Extraction and precipitation were repeated as before. The mRNA was then isolated with the Qiagen Oligotex kit (QIAGEN, Inc., Chatsworth, CA) and used to construct the cDNA libraries.

The mRNAs were handled according to the recommended protocols in the SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning (Cat. #18248-013, Gibco/BRL, Gaithersburg, MD). cDNAs were fractionated on a Sepharose CL4B column (Cat. #275105-01, Pharmacia), and those cDNAs exceeding 400 bp were ligated into pINCY I. The plasmid pINCY I was subsequently transformed into DH5a™ competent cells (Cat. #18258-012, Gibco/BRL).

II Isolation and Sequencing of cDNA Clones

Plasmid DNA was released from the cells and purified using the REAL Prep 96 Plasmid Kit (Catalog #26173, QIAGEN, Inc.). This kit enabled the simultaneous purification of 96 samples in a 96-well block using multi-channel reagent dispensers. The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Catalog #22711, Gibco/BRL) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) after inoculation, the cultures were incubated for 19 hours and at the end of incubation, the cells were lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the plasmid DNA pellet was resuspended in 0.1 ml of distilled water. After the last step in the protocol, samples were transferred to a 96-well block for storage at 4° C.

The cDNAs were sequenced by the method of Sanger et al. (1975, J. Mol. Biol. 94:441f),

BNSDOCID <WO___9840494A1_I_>

using a Hamilton Micro Lab 2200 (Hamilton, Reno, NV) in combination with Peltier Thermal Cyclers (PTC200 from MJ Research, Watertown, MA) and Applied Biosystems 377 DNA Sequencing Systems; and the reading frame was determined.

III Homology Searching of cDNA Clones and Their Deduced Proteins

The nucleotide sequences of the Sequence Listing or amino acid sequences deduced from them were used as query sequences against databases such as GenBank, SwissProt, BLOCKS, and Pima II. These databases which contain previously identified and annotated sequences were searched for regions of homology (similarity) using BLAST, which stands for Basic Local Alignment Search Tool (Altschul (1993) supra, Altschul (1990) supra).

BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs which may be of prokaryotic (bacterial) or eukaryotic (animal, fungal or plant) origin. Other algorithms such as the one described in Smith R.F. and T.F. Smith (1992, Protein Engineering 5:35-51), incorporated herein by reference, can be used when dealing with primary sequence patterns and secondary structure gap penalties. As disclosed in this application, the sequences have lengths of at least 49 nucleotides, and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

The BLAST approach, as detailed in Karlin and Altschul (<u>supra</u>) and incorporated herein by reference, searches for matches between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. In this application, threshold was set at 10^{-25} for nucleotides and 10^{-14} for peptides.

Incyte nucleotide sequences were searched against the GenBank databases for primate (pri), rodent (rod), and mammalian sequences (mam), and deduced amino acid sequences from the same clones are searched against GenBank functional protein databases, mammalian (mamp), vertebrate (vrtp) and eukaryote (eukp), for homology. The relevant database for a particular match were reported as a GIxxx±p (where xxx is pri, rod, etc and if present, p = peptide). The product score is calculated as follows: the % nucleotide or amino acid identity [between the query and reference sequences] in BLAST is multiplied by the % maximum possible BLAST score [based on the lengths of query and reference sequences] and then divided by 100. Where an Incyte Clone was homologous to several sequences, up to five matches were provided with their relevant scores. In an analogy to the hybridization procedures used in the laboratory, the electronic stringency for an exact match was set at 70, and the conservative lower limit for an

BNSDOCID: <WO___9840494A1_I_>

exact match was set at approximately 40 (with 1-2% error due to uncalled bases).

IV Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook et al., supra).

Analogous computer techniques using BLAST (Altschul, S.F. 1993 and 1990, supra) are used to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ™ database (Incyte Pharmaceuticals). This analysis is much faster than multiple, membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or homologous.

The basis of the search is the product score which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analysis are reported as a list of libraries in which the transcript encoding REE-2 occurs. Abundance and percent abundance are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and percent abundance is abundance divided by the total number of sequences examined in the cDNA library.

V Extension of REE-2-Encoding Polynucleotides

Nucleic acid sequence of Incyte clone 1646823 or SEQ ID NO:2 is used to design oligonucleotide primers for extending a partial nucleotide sequence to full length or for obtaining 5' or 3', intron or other control sequences from genomic libraries. One primer is synthesized to initiate extension in the antisense direction (XLR) and the other is synthesized to extend sequence in the sense direction (XLF). Primers are used to facilitate the extension of the known sequence "outward" generating amplicons containing new, unknown nucleotide sequence for the region of interest. The initial primers are designed from the cDNA using OLIGO 4.06 (National Biosciences), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C.

BNSDOCID: <WO___9840494A1_I_>

Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations is avoided.

The original, selected cDNA libraries, or a human genomic library are used to extend the sequence; the latter is most useful to obtain 5' upstream regions. If more extension is necessary or desired, additional sets of primers are designed to further extend the known region.

By following the instructions for the XL-PCR kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix, high fidelity amplification is obtained. Beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, PCR is performed using the Peltier Thermal Cycler (PTC200; M.J. Research, Watertown, MA) and the following parameters:

Step 1	94° C for 1 min (initial denaturation)
Step 2	65° C for 1 min
Step 3	68° C for 6 min
Step 4	94° C for 15 sec
Step 5	65° C for 1 min
Step 6	68° C for 7 min
Step 7	Repeat step 4-6 for 15 additional cycles
Step 8	94° C for 15 sec
Step 9	65° C for 1 min
Step 10	68° C for 7:15 min
Step 11	Repeat step 8-10 for 12 cycles
Step 12	72° C for 8 min
Step 13	4° C (and holding)

A 5-10 μl aliquot of the reaction mixture is analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products are selected and removed from the gel. Further purification involves using a commercial gel extraction method such as QIAQuick™ (QIAGEN Inc., Chatsworth, CA). After recovery of the DNA, Klenow enzyme is used to trim single-stranded, nucleotide overhangs creating blunt ends which facilitate religation and cloning.

After ethanol precipitation, the products are redissolved in 13 μl of ligation buffer, 1μl T4-DNA ligase (15 units) and 1μl T4 polynucleotide kinase are added, and the mixture is incubated at room temperature for 2-3 hours or overnight at 16° C. Competent E. coli cells (in 40 μl of appropriate media) are transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium (Sambrook et al., supra). After incubation for one hour at 37° C, the whole transformation mixture is plated on Luria Bertani (LB)-agar (Sambrook et al., supra) containing

BNSDOCID <WO__9840494A1_I_>

2x Carb. The following day, several colonies are randomly picked from each plate and cultured in 150 μl of liquid LB/2x Carb medium placed in an individual well of an appropriate, commercially-available, sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture is transferred into a non-sterile 96-well plate and after dilution 1:10 with water, 5 μl of each sample is transferred into a PCR array.

For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for the extension reaction are added to each well. Amplification is performed using the following conditions:

Step 1	94° C for 60 sec
Step 2	94° C for 20 sec
Step 3	55° C for 30 sec
Step 4	72° C for 90 sec
Step 5	Repeat steps 2-4 for an additional 29 cycles
Step 6	72° C for 180 sec
Step 7	4° C (and holding)

Aliquots of the PCR reactions are run on agarose gels together with molecular weight markers. The sizes of the PCR products are compared to the original partial cDNAs, and appropriate clones are selected, ligated into plasmid, and sequenced.

VI Labeling and Use of Hybridization Probes

Hybridization probes derived from SEQ ID NO:2 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 (National Biosciences), labeled by combining 50 pmol of each oligomer and 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham) and T4 polynucleotide kinase (DuPont NEN®, Boston, MA). The labeled oligonucleotides are substantially purified with Sephadex G-25 superfine resin column (Pharmacia & Upjohn). A portion containing 10^7 counts per minute of each of the sense and antisense oligonucleotides is used in a typical membrane based hybridization analysis of human genomic DNA digested with one of the following endonucleases (Ase I, Bgl II, Eco RI, Pst I, Xba 1, or Pvu II; DuPont NEN®).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham, NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5%

BNSDOCID <WO___9840494A1_I_>

sodium dodecyl sulfate. After XOMAT AR™ film (Kodak, Rochester, NY) is exposed to the blots in a Phosphoimager cassette (Molecular Dynamics, Sunnyvale, CA) for several hours, hybridization patterns are compared visually.

VII Antisense Molecules

Antisense molecules or the complement of the REE-2-encoding sequence, or any part thereof, is used to inhibit in vivo or in vitro expression of naturally occurring REE-2. Although use of antisense oligonucleotides, comprising about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. An oligonucleotide based on the coding sequences of REE-2, as shown in Figures 1A and 1B, is used to inhibit expression of naturally occurring REE-2. The complementary oligonucleotide is designed from the most unique 5' sequence as shown in Figures 1A and 1B and used either to inhibit transcription by preventing promoter binding to the upstream nontranslated sequence or translation of an REE-2-encoding transcript by preventing the ribosome from binding. Using an appropriate portion of the signal and 5' sequence of SEQ ID NO:2, an effective antisense oligonucleotide includes any 15-20 nucleotides spanning the region which translates into the signal or 5' coding sequence of the polypeptide as shown in Figures 1A and 1B.

VIII Expression of REE-2

Expression of REE-2 is accomplished by subcloning the cDNAs into appropriate vectors and transforming the vectors into host cells. In this case, the cloning vector, pSport, previously used for the generation of the cDNA library is used to express REE-2 in E. coli. Upstream of the cloning site, this vector contains a promoter for ß-galactosidase, followed by sequence containing the amino-terminal Met, and the subsequent seven residues of ß-galactosidase. Immediately following these eight residues is a bacteriophage promoter useful for transcription and a linker containing a number of unique restriction sites.

Induction of an isolated, transformed bacterial strain with IPTG using standard methods produces a fusion protein which consists of the first eight residues of ß-galactosidase, about 5 to 15 residues of linker, and the full length protein. The signal residues direct the secretion of REE-2 into the bacterial growth media which can be used directly in the following assay for activity.

IX Demonstration of REE-2 Activity

REE's deaminase activity can be measured by a method described by MacGinntie A.J. et al. (1995, J. Biol. Chem. 270: 14768-14775). Substantially purified REE is incubated with 3.3 uCi of [^{3}H] deoxycytidine and 250 uM cytidine in a total volume of 10 ul in a buffer containing 45 mM TRIS, pH 7.5. After timed incubations the reaction is quenched by the addition of 2 ul of

BNSDOCID <WO__9840494A1_I_>

10 ug/ul each deoxycytidine and deoxyuridine. Any insoluble material is removed by centrifugation for 2 minutes at full speed in a microcentrifuge, and 4 ul of the reaction mixture is applied to a polyethyleneimine-cellulose thin layer chromatographic plate. The corresponding deoxycytidine and deoxyuridine bands are visualized by exposure to UV light and scraped into scintillation fluid for quantification by liquid scintillation spectroscopy.

X Production of REE-2 Specific Antibodies

REE-2 that is substantially purified using PAGE electrophoresis (Sambrook, supra), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols. The amino acid sequence deduced from SEQ ID NO:2 is analyzed using DNASTAR software (DNASTAR Inc) to determine regions of high immunogenicity and a corresponding oligopolypeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions, is described by Ausubel et al. (supra), and others.

Typically, the oligopeptides are 15 residues in length, synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry, and coupled to keyhole limpet hemocyanin (KLH, Sigma, St. Louis, MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel et al., supra). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity, for example, by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated, goat anti-rabbit IgG.

XI Purification of Naturally Occurring REE-2 Using Specific Antibodies

Naturally occurring or recombinant REE-2 is substantially purified by immunoaffinity chromatography using antibodies specific for REE-2. An immunoaffinity column is constructed by covalently coupling REE-2 antibody to an activated chromatographic resin, such as CnBr-activated Sepharose (Pharmacia & Upjohn). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing REE-2 is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of REE-2 (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/REE-2 binding (eg, a buffer of pH 2-3 or a high concentration of a chaotrope, such as urea or thiocyanate ion), and REE-2 is collected.

XII Identification of Molecules Which Interact with REE-2

REE-2 or biologically active fragments thereof are labeled with ^{125}I Bolton-Hunter

BNSDOCID: <WO___9840494A1_I_>

reagent (Bolton et al. (1973) Biochem. J. 133: 529). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled REE-2, washed and any wells with labeled REE-2 complex are assayed. Data obtained using different concentrations of REE-2 are used to calculate values for the number, affinity, and association of REE-2 with the candidate molecules.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

BNSDOCID: <WO___9840494A1_I_>

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: INCYTE PHARMACEUTICALS, INC.

(ii) TITLE OF THE INVENTION: NOVEL RNA EDITING ENZYME

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) PCT APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Herewith

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/816,241
(B) FILING DATE: 13-MAR-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0239 PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 650-855-0555
(B) TELEFAX: 650-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 190 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: PROSTUT09
(B) CLONE: 1646823

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Asn Pro Gln Ile Arg Asn Pro Met Lys Ala Met Tyr Pro Gly Thr
 1               5                  10                  15
Phe Tyr Phe Gln Phe Lys Asn Leu Trp Glu Ala Asn Asp Arg Asn Glu
            20                  25                  30
Thr Trp Leu Cys Phe Thr Val Glu Gly Ile Lys Arg Arg Ser Val Val
        35                  40                  45
Ser Trp Lys Thr Gly Val Phe Arg Asn Gln Val Asp Ser Glu Thr His
    50                  55                  60
Cys His Ala Glu Arg Cys Phe Leu Ser Trp Phe Cys Asp Asp Ile Leu
65                  70                  75                  80
```


BNSDOCID <WO__9840494A1_I_>

```
Ser Pro Asn Thr Lys Tyr Gln Val Thr Trp Tyr Thr Ser Trp Ser Pro
                85                  90                  95
Cys Pro Asp Cys Ala Gly Glu Val Ala Glu Phe Leu Ala Arg His Ser
            100                 105                 110
Asn Val Asn Leu Thr Ile Phe Thr Ala Arg Leu Tyr Tyr Phe Gln Tyr
        115                 120                 125
Pro Cys Tyr Gln Glu Gly Leu Arg Ser Leu Ser Gln Glu Gly Val Ala
    130                 135                 140
Val Glu Ile Met Asp Tyr Glu Asp Phe Lys Tyr Cys Trp Glu Asn Phe
145                 150                 155                 160
Val Tyr Asn Asp Asn Glu Pro Phe Lys Pro Trp Lys Gly Leu Lys Thr
                165                 170                 175
Asn Phe Arg Leu Leu Lys Arg Arg Leu Arg Glu Ser Leu Gln
            180                 185                 190
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 610 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: PROSTUT09
(B) CLONE: 1646833

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
ATGAATCCAC AGATCAGAAA CCCGATGAAG GCAATGTATC CAGGCACATT CTACTTCCAA     60
TTTAAAAACC TATGGGAAGC CAACGATCGG AACGAAACTT GGCTGTGCTT CACCGTGGAA    120
GGTATAAAGC GCCGCTCAGT TGTCTCCTGG AAGACGGGCG TCTTCCGAAA CCAGGTGGAT    180
TCTGAGACCC ATTGTCATGC AGAAAGGTGC TTCCTCTCTT GGTTCTGCGA CGACATACTG    240
TCTCCTAACA CAAAGTACCA GGTCACCTGG TACACATCTT GGAGCCCTTG CCCAGACTGT    300
GCAGGGGAGG TGGCCGAGTT CCTGGCCAGG CACAGCAACG TGAATCTCAC CATCTTCACC    360
GCCCGCCTCT ACTACTTCCA GTATCCATGT TACCAGGAGG GGCTCCGCAG CCTGAGTCAG    420
GAAGGGGTCG CTGTGGAGAT CATGGACTAT GAAGATTTTA AATATTGTTG GGAAAACTTT    480
GTGTACAATG ATAATGAGCC ATTCAAGCCT TGGAAGGGAT TAAAAACCAA CTTTCGACTT    540
CTGAAAAGAA GGCTACGGGA GAGTCTCCAG TGAGGGGTCT CCCTGGGCCT CATGGTCTGT    600
CTCCTCTAAG                                                           610
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 116 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 436941

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Asn Ser Ala Arg Glu Ile Tyr Arg Val Thr Trp Phe Ile Ser Trp Ser
 1               5                  10                  15
Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala Phe Leu Gln
            20                  25                  30
Glu Asn Thr His Val Arg Leu Pro Ile Phe Ala Ala Arg Ile Tyr Asp
        35                  40                  45
Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg Asp Ala Gly
    50                  55                  60
Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Glu Tyr Cys Trp Asp
65                  70                  75                  80
```


BNSDOCID <WO__9840494A1_I_>

```
Thr Phe Val Tyr Arg Gln Gly Cys Pro Phe Gln Pro Trp Asp Gly Leu
                85                  90                  95
Glu Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala Ile Leu Gln
            100                 105                 110
Asn Gln Gly Asn
        115
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 236 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 11777906

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Thr Ser Glu Lys Gly Pro Ser Thr Gly Asp Pro Thr Leu Arg Arg
 1               5                  10                  15
Arg Ile Glu Pro Trp Glu Phe Asp Val Phe Tyr Asp Pro Arg Glu Leu
            20                  25                  30
Arg Lys Glu Ala Cys Leu Leu Tyr Glu Ile Lys Trp Gly Met Ser Arg
        35                  40                  45
Lys Ile Trp Arg Ser Ser Gly Lys Asn Thr Thr Asn His Val Glu Val
    50                  55                  60
Asn Phe Ile Lys Lys Phe Thr Ser Glu Arg Asp Phe His Pro Ser Ile
65                  70                  75                  80
Ser Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Trp Glu Cys
                85                  90                  95
Ser Gln Ala Ile Arg Glu Phe Leu Ser Arg His Pro Gly Val Thr Leu
            100                 105                 110
Val Ile Tyr Val Ala Arg Leu Phe Trp His Met Asp Gln Gln Asn Arg
        115                 120                 125
Gln Gly Leu Arg Asp Leu Val Asn Ser Gly Val Thr Ile Gln Ile Met
    130                 135                 140
Arg Ala Ser Glu Tyr Tyr His Cys Trp Arg Asn Phe Val Asn Tyr Pro
145                 150                 155                 160
Pro Gly Asp Glu Ala His Trp Pro Gln Tyr Pro Pro Leu Trp Met Met
                165                 170                 175
Leu Tyr Ala Leu Glu Leu His Cys Ile Ile Leu Ser Leu Pro Pro Cys
            180                 185                 190
Leu Lys Ile Ser Arg Arg Trp Gln Asn His Leu Thr Phe Phe Arg Leu
        195                 200                 205
His Leu Gln Asn Cys His Tyr Gln Thr Ile Pro Pro His Ile Leu Leu
    210                 215                 220
Ala Thr Gly Leu Ile His Pro Ser Val Ala Trp Arg
225                 230                 235
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 229 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 585813

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

BNSDOCID <WO__9840494A1_I_>

Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
1 5 10 15

Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu
20 25 30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly Arg His
35 40 45

Ser Ile Trp Arg His Thr Ser Gln Asn Thr Asn Lys His Val Glu Val
50 55 60

Asn Phe Ile Glu Lys Phe Thr Thr Glu Arg Tyr Phe Cys Pro Asn Thr
65 70 75 80

Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys
85 90 95

Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg Tyr Pro His Val Thr Leu
100 105 110

Phe Ile Tyr Ile Ala Arg Leu Tyr His His Ala Asp Pro Arg Asn Arg
115 120 125

Gln Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln Ile Met
130 135 140

Thr Glu Gln Glu Ser Gly Tyr Cys Trp Arg Asn Phe Val Asn Tyr Ser
145 150 155 160

Pro Ser Asn Glu Ala His Trp Pro Arg Tyr Pro His Leu Trp Val Arg
165 170 175

Leu Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys
180 185 190

Leu Asn Ile Leu Arg Arg Lys Gln Pro Gln Leu Thr Phe Phe Thr Ile
195 200 205

Ala Leu Gln Ser Cys His Tyr Gln Arg Leu Pro Pro His Ile Leu Trp
210 215 220

Ala Thr Gly Leu Lys
225

BNSDOCID <WO___9840494A1_I_>

What is claimed is:

1. A substantially purified human RNA editing enzyme comprising the amino acid sequence of SEQ ID NO:1 or fragments thereof.

2. An isolated and purified polynucleotide sequence encoding the human RNA editing enzyme of claim 1.

3. A polynucleotide sequence which hybridizes under stringent conditions to the polynucleotide sequence of claim 2.

4. A hybridization probe comprising the polynucleotide sequence of claim 2.

5. An isolated and purified polynucleotide sequence comprising SEQ ID NO:2 or variants thereof.

6. A polynucleotide sequence which is complementary to the polynucleotide sequence of claim 2 or variants thereof.

7. A hybridization probe comprising the polynucleotide sequence of claim 6.

8. An expression vector containing the polynucleotide sequence of claim 2.

9. A host cell containing the vector of claim 8.

10. A method for producing a polypeptide comprising the amino acid sequence of SEQ ID NO:1 the method comprising the steps of:

a) culturing the host cell of claim 9 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

11. A pharmaceutical composition comprising a substantially purified human RNA editing enzyme having an amino acid sequence of claim 1 in conjunction with a suitable pharmaceutical carrier.

12. A purified antibody which binds specifically to the polypeptide of claim 1.

13. A purified agonist which specifically modulates the activity of the polypeptide of claim 1.

14. A purified antagonist which specifically modulates the activity of the polypeptide of claim 1.

15. A method for treating cancer comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 11.

16. A method for treating viral diseases comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 11.

BNSDOCID <WO__9840494A1_I_>

17. A method for treating an autoimmune disorder comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 14.

18. A method for treating a circulatory system disorder comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 11.

19. A method for treating a neurological disease comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 14.

20. A method for detection of a polynucleotide which encodes human RNA editing enzyme in a biological sample comprising the steps of:

a) hybridizing the polynucleotide of claim 6 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and

b) detecting said hybridization complex, wherein the presence of said complex correlates with the presence of a polynucleotide encoding human RNA editing enzyme in said biological sample.

BNSDOCID <WO___9840494A1_I_>

```
        9          18          27          36          45          54
ATG AAT CCA CAG ATC AGA AAC CCG ATG AAG GCA ATG TAT CCA GGC ACA TTC TAC
M   N   P   Q   I   R   N   P   M   K   A   M   Y   P   G   T   F   Y

       63          72          81          90          99         108
TTC CAA TTT AAA AAC CTA TGG GAA GCC AAC GAT CGG AAC GAA ACT TGG CTG TGC
F   Q   F   K   N   L   W   E   A   N   D   R   N   E   T   W   L   C

      117         126         135         144         153         162
TTC ACC GTG GAA GGT ATA AAG CGC CGC TCA GTT GTC TCC TGG AAG ACG GGC GTC
F   T   V   E   G   I   K   R   R   S   V   V   S   W   K   T   G   V

      171         180         189         198         207         216
TTC CGA AAC CAG GTG GAT TCT GAG ACC CAT TGT CAT GCA GAA AGG TGC TTC CTC
F   R   N   Q   V   D   S   E   T   H   C   H   A   E   R   C   F   L

      225         234         243         252         261         270
TCT TGG TTC TGC GAC GAC ATA CTG TCT CCT AAC ACA AAG TAC CAG GTC ACC TGG
S   W   F   C   D   D   I   L   S   P   N   T   K   Y   Q   V   T   W

      279         288         297         306         315         324
TAC ACA TCT TGG AGC CCT TGC CCA GAC TGT GCA GGG GAG GTG GCC GAG TTC CTG
Y   T   S   W   S   P   C   P   D   C   A   G   E   V   A   E   F   L

      333         342         351         360         369         378
GCC AGG CAC AGC AAC GTG AAT CTC ACC ATC TTC ACC GCC CGC CTC TAC TAC TTC
A   R   H   S   N   V   N   L   T   I   F   T   A   R   L   Y   Y   F
```

FIGURE 1A

```
                387         396         405         414         423         432
CAG TAT CCA TGT TAC CAG GAG GGG CTC CGC AGC CTG AGT CAG GAA GGG GTC GCT
Q   Y   P   C   Y   Q   E   G   L   R   S   L   S   Q   E   G   V   A

                441         450         459         468         477         486
GTG GAG ATC ATG GAC TAT GAA GAT TTT AAA TAT TGT TGG GAA AAC TTT GTG TAC
V   E   I   M   D   Y   E   D   F   K   Y   C   W   E   N   F   V   Y

                495         504         513         522         531         540
AAT GAT AAT GAG CCA TTC AAG CCT TGG AAG GGA TTA AAA ACC AAC TTT CGA CTT
N   D   N   E   P   F   K   P   W   K   G   L   K   T   N   F   R   L

                549         558         567         576         585         594
CTG AAA AGA AGG CTA CGG GAG AGT CTC CAG TGA GGG GTC TCC CTG GGC CTC ATG
L   K   R   R   L   R   E   S   L   Q

                603
GTC TGT CTC CTC TAA G
```

FIGURE 1B

```
1    M N P Q I - - - - - - R N P M K A M Y P G T F - - Y F Q F K   SEQ ID NO-1
1    N S A - - - - - - - - - - - - - - - - - - - - - - - - - - -   436941
1    M T S E K G P S T G D P T L R R R I E P W E F D V F Y D P R   GI 1177798
1    M S S E T G P V A V D P T L R R R I E P H E F E V F F D P R   GI 585813

23   N L W E A N D R N E T W L C F T V E G I K R R S V V S W K T   SEQ ID NO-1
4    - - - - - - - - - - - - - - - - - - - - - - R E I - - - - -   436941
31   E L - - - - - R K E A C L L Y E I K W G M S R K I - - W R S   GI 1177798
31   E L - - - - - R K E T C L L Y E I N W G G R H S I - - W R H   GI 585813

53   G V F R N Q V D S E T H C H A E R C F L S W F C D D - I L S   SEQ ID NO-1
7    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   436941
54   S - - - - - - G K N T T N H V E V N F I K K F T S E R D F H   GI 1177798
54   T - - - - - - S Q N T N K H V E V N F I E K F T T E R Y F C   GI 585813

82   P N T K Y Q V T W Y T S W S P C P - - D C A G E V A E F L A   SEQ ID NO-1
7    - - - - Y R V T W F I S W S P C F S W G C A G E V R A F L Q   436941
78   P S I S C S I T W F L S W S P C - - W E C S Q A I R E F L S   GI 1177798
78   P N T R C S I T W F L S W S P C - - G E C S R A I T E F L S   GI 585813

110  R H S N V N L T I F T A R L Y Y F Q Y P C Y Q E G L R S L S   SEQ ID NO-1
33   E N T H V R L P I F A A R I Y D Y D - P L Y K E A L Q M L R   436941
106  R H P G V T L V I Y V A R L F W H M D Q Q N R Q G L R D L V   GI 1177798
106  R Y P H V T L F I Y I A R L Y H H A D P R N R Q G L R D L I   GI 585813
```

FIGURE 2A

```
140 Q E G V A V E I M D Y E D F K Y C W E N F V - Y N D N E - -  SEQ ID NO-1
62  D A G A Q V S I M T Y D E F E Y C W D T F V - Y R Q G C - -  436941
136 N S G V T I Q I M R A S E Y Y H C W R N F V N Y P P G D E A  GI 1177798
136 S S G V T I Q I M T E Q E S G Y C W R N F V N Y S P S N E A  GI 585813

167 - - - P F K P - W K - - - - - - - - - - - - G L K T N F R L  SEQ ID NO-1
89  - - - P F Q P - W D - - - - - - - - - - - - G L E E H S Q A  436941
166 H W P Q Y P P L W M M L Y A L E L H C I I L S L P P C L K I  GI 1177798
166 H W P R Y P H L W V R L Y V L E L Y C I I L G L P P C L N I  GI 585813

181 L K R R - - - - - - L R E S L Q                              SEQ ID NO-1
103 L S G R - - - - - - L R A I L Q N - - - - - - - - - - - - -  436941
196 S R R W Q N H L T F F R L H L Q N C H Y Q T I P P H I L L A  GI 1177798
196 L R R K Q P Q L T F F T I A L Q S C H Y Q R L P P H I L W A  GI 585813
 exactly.
190                                                              SEQ ID NO-1
114 Q G - - - - - - - - N                                        436941
226 T G L I H P S V A W R                                        GI 1177798
226 T G L - - - - - - - K                                        GI 585813
```

FIGURE 2B

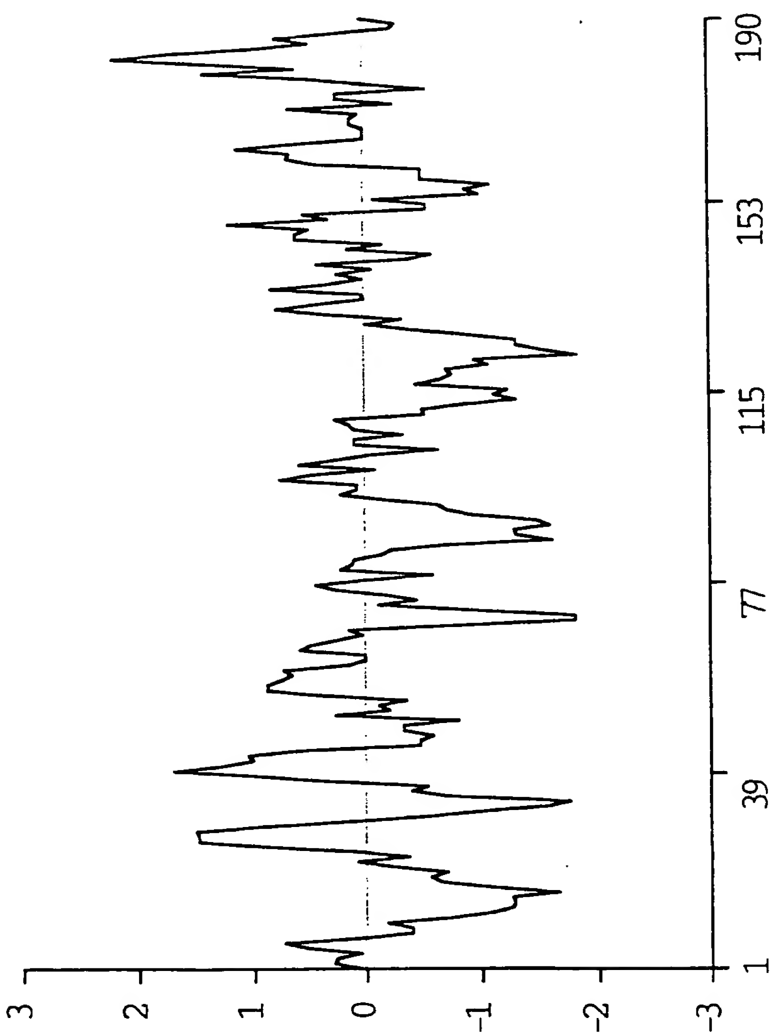


FIGURE 3

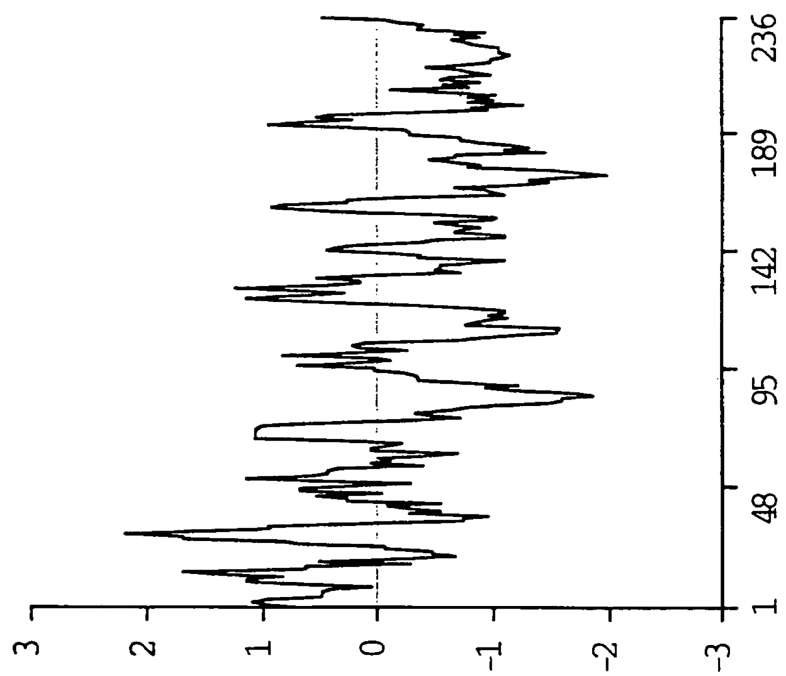


FIGURE 4

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/04565

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/52 C12N9/00 A61K38/43 C07K16/40 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A61K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 18316 A (ARCH DEV CORP) 18 August 1994	1-11
A	Compare amino acids 93-97 in SEQ ID NO:1 of the present application, with amino acids 89-93 in SEQ ID NO:2, on page 65 of WO 94/18316. see page 49, line 24 - line 30 ---	12-14
X	Hillier, L et al.,:The WashU-Merck EST project-Seq HS1150939 Accession number AA2233019; 6 March 1997. zr46a07r1 Soares NhHMPuS1 Homo sapiens cDNA clone 666420 5' XP002071164 Compare nucleotides 1-123 with nucleotides 546-608 in SEQ ID NO:2 of the present application. ---	3-7

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 July 1998	28.07.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Cupido, M

2

INTERNATIONAL SEARCH REPORT

Int tional Application No
PCT/US 98/04565

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WOOLF T M ET AL: "Toward the therapeutic editing of mutated RNA sequences" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 92, 1 August 1995, pages 8298-8302, XP000574995 see page 8301 discussion ---	11,15-19
P,X	WO 98 04714 A (HAWKINS PHILLIP R ;INCYTE PHARMA INC (US); AU YOUNG JANICE (US); H) 5 February 1998 Compare amino acids 88-91 in SEQ ID NO:1 of the present application, with amino acids 119-122 in SEQ ID NO:1, on page 40 of WO 98/04714. -----	1-11

2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 98/04565

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos
because they relate to subject matter not required to be searched by this Authority, namely:
Although claims 15-19 are directed to a method of treatment of the human or animal body, the search has been carried out and based on the alleged effects of the composition.

2. [ ] Claims Nos
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos..
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 98/04565

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9418316 A	18-08-1994	US 5550034 A	27-08-1996
		US 5434058 A	18-07-1995
		AU 6136094 A	29-08-1994
		EP 0682701 A	22-11-1995
WO 9804714 A	05-02-1998	US 5747319 A	05-05-1998
		AU 3741997 A	20-02-1998